# Quantifying the sequence–function relation in gene silencing by bacterial small RNAs

Yue Hao[a,1], Zhongge J. Zhang[b,1], David W. Erickson[c,d], Min Huang[a], Yingwu Huang[e], Junbai Li[f], Terence Hwa[b,c,d,2], and Hualin Shi[a,2]

[a]Institute of Theoretical Physics and [f]Institute of Chemistry, Chinese Academy of Sciences, Beijing 100190, China; [b]Section of Molecular Biology, Division of Biological Sciences, [c]Center for Theoretical Biological Physics, and [d]Department of Physics, University of California at San Diego, La Jolla, CA 92093-0374; and [e]Department of Biological Sciences and Biotechnology, Tsinghua University, Beijing 100084, China

Edited* by Bonnie L. Bassler, Howard Hughes Medical Institute and Princeton University, Princeton, NJ, and approved June 1, 2011 (received for review January 13, 2011)

**Sequence–function relations for small RNA (sRNA)-mediated gene silencing were quantified for the sRNA RyhB and some of its mRNA targets in *Escherichia coli*. Numerous mutants of RyhB and its targets were generated and their in vivo functions characterized at various levels of target and RyhB expression. Although a core complementary region is required for repression by RyhB, variations in the complementary sequences of the core region gave rise to a continuum of repression strengths, correlated exponentially with the computed free energy of RyhB-target duplex formation. Moreover, sequence variations in the linker region known to interact with the RNA chaperone Hfq also gave rise to a continuum of repression strengths, correlated exponentially with the computed energy cost of keeping the linker region open. These results support the applicability of the thermodynamic model in predicting sRNA–mRNA interaction and suggest that sequences at these locations may be used to fine-tune the degree of repression. Surprisingly, a truncated RyhB without the Hfq-binding region is found to repress multiple targets of the wild-type RyhB effectively, both in the presence and absence of Hfq, even though the former is required for the activity of wild-type RyhB itself. These findings challenge the commonly accepted model concerning the function of Hfq in gene silencing—both in providing stability to the sRNAs and in catalyzing the target mRNAs to take on active conformations—and raise the intriguing question of why many endogenous sRNAs subject their functions to Hfq-dependences.**

gene regulation | noncoding RNA | posttranscriptional control | quantitative biology | RNA interaction

A significant development in gene regulation in the past decade is a growing appreciation of the complex roles that small regulatory RNA (sRNA) can play in coordinating gene activities in both prokaryotes and eukaryotes (1–3). In *Escherichia coli*, approximately 80 sRNA genes have been identified (3). There exists by now a basic understanding of the molecular components and mechanisms involved, at least for a major class of bacterial sRNA that acts *in trans* through base pairing (4–15). Recent theoretical and experimental studies have further revealed unique functional features of sRNA-mediated gene regulation (9, 16–20): because of the stoichiometric mode of target inactivation, sRNA-mediated regulation exhibits an abrupt and sensitive response to input signals while being robust to stochastic fluctuations.

How is this mode of regulation encoded in the molecular sequences of the sRNA and its targets? In the case of transcriptional regulation, a great deal is known quantitatively about the interaction between a DNA binding sequence (operator) and its cognate transcription factor (TF) and the regulatory consequences of this interaction: similarity of the operator to its "consensus sequence" determines its binding affinity to the cognate TF (21–24), and the latter in turn affects the rate of transcriptional initiation (25). Such knowledge, obtained by quantitative experimental studies of a few exemplary TFs decades ago (21–23), led to the later development of powerful bioinformatic approaches for the discovery of TF binding sites from genomic analysis (26), quantitative analysis of transcriptional regulation for complex promoters and even realistic modeling of promoter evolution (27, 28). Knowledge of sequence–function relation for sRNA-mediated gene regulation could lead to similar progress in bioinformatic identification of sRNA genes and their targets, in quantitative modeling of sRNA-mediated genetic circuits and their evolution.

Such sequence–function relation has been characterized to some extent for the best-studied class of bacterial sRNAs involving the RNA chaperon Hfq (1, 29). Members of this class include OxyS (4, 30), DsrA (5, 31), RyhB (6, 32), Spot42 (7), SgrS (8, 13), MicC (10), MicA (11), and MicF (12). Common structural features of these sRNAs include one or more target interaction regions, each contained in a hairpin loop, an unstructured Hfq-binding linker region, and a Rho-independent terminator at the 3′ end (1). Hfq binds to many mRNAs and sRNAs (33). It is known to protect some sRNAs and mRNAs from rapid degradation (11, 32, 34) and stimulate the interaction between various sRNAs and their targets (13, 30, 35, 36). Much work has gone into defining the interaction region of the sRNA, usually complementary to the translation initiation regions of the targets (1, 3, 29), although pairings in the coding sequence have also been reported recently (37, 38). Extensive in vivo characterization of sequence–function relation was performed on the repression of *ptsG* mRNA by the sRNA SgrS in response to sugar phosphate stress (13, 39). Scanning by single base substitution throughout the interaction region, a core interaction region consisting of six bases was identified to be required for exerting repression function (13), whereas replacement of bases flanking the core hardly affected repression (39). Similar results were echoed by studies in other systems (5, 10, 37, 38). Compared with the interaction region, the Hfq binding region has not been as extensively characterized, other than its preference for AU-rich sequences (30, 40).

Among the existing studies characterizing the sequence–function relation, most have been done at a qualitative level (i.e., whether a certain sRNA sequence repressed its targets). On the other hand, the threshold-linear response of sRNA-mediated regulation depends quantitatively on the energetics and kinetics of the sRNA-target interaction (16, 18), specified through the RNA sequences. Quantitative knowledge of the sequence–function relation may therefore lead to another layer of appreciation of how the sRNA systems work in vivo.

In this study, we took a first step toward quantifying the in vivo sequence–function relation, focusing on the regulation of *sodB* expression by the sRNA RyhB in *E. coli*. RyhB, expressed under low $Fe^{2+}$ conditions and central to the iron homeostasis con-

Author contributions: Y. Huang, H.S., and T.H. designed research; Y. Hao, Z.J.Z., D.W.E., M.H., Y. Huang, T.H., and H.S. performed research; Y. Hao, Z.J.Z., M.H., Y. Huang, and J.L. contributed new reagents/analytic tools; Y. Hao, Z.J.Z., D.W.E., M.H., T.H., and H.S. analyzed data; and Y. Hao, Z.J.Z., D.W.E., T.H., and H.S. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]Y. Hao and Z.J.Z. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: shihl@itp.ac.cn or hwa@ucsd.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1100432108/-/DCSupplemental.





MICROBIOLOGY

trol (41), is one of the best-characterized members of the Hfq-dependent class of sRNA (1). *sodB* mRNA, encoding a superoxide dismutase expressed under the condition of high $Fe^{2+}$ level, is one of the most prominent targets of RyhB (32, 41). To quantify the sequence determinants of RyhB–*sodB* interaction, we generated a large number of targeted mutants in both the interaction region and the Hfq-binding region and characterized the interactions of the mutants quantitatively using translational *sodB-gfp* fusion constructs, expressed under the control of a titratable promoter (16). The sequence–function relations obtained were then further correlated with the energetics of RyhB–*sodB* interaction through RNA secondary structure analysis. A plethora of results were obtained, including the graded nature of RyhB–*sodB* interaction, tunable by the composition of complementary sequences in the core interaction region and by the composition of the Hfq-binding region. Surprisingly, a truncated RyhB without the Hfq-binding region is found to be sufficient to repress targets of the wild-type RyhB, both in the presence and absence of Hfq. Implications of these results are discussed.

## Results

RyhB has a distinct secondary structure consisting of three hairpin loops with the Hfq-binding linker (Fig. S1, *Upper*). The loop regions mediate interaction with different mRNA targets (1); Fig. S1 shows an example of the core contacts (red nucleotides) between RyhB and its most prominent target, the 5′ region of the *sodB* mRNA (32, 41).

To elucidate the sequence determinant of RyhB–*sodB* interaction, we generated mutations of RyhB and *sodB* in their respective control regions (*SI Materials and Methods*) and examined their effects on the repression of *sodB* expression using a strategy introduced earlier by Levine et al. (16). To quantify the degree of repression, we fused the 5′ UTR of each mutated *sodB* along with its first 11 codons, including the entire control region of *sodB* mRNA (Fig. S1), to the 5′ end of the *gfp* structural gene. The resulting fusion gene was incorporated into the pZE12S-series plasmid under the control of a synthetic $P_{Llac\text{-}O1}$ promoter inducible by isopropyl β-D-thiogalactoside (IPTG). Analogously, each mutated *ryhB* gene was incorporated into the pZA31R-series plasmid under the control of a synthetic $P_{LTet\text{-}O1}$ promoter inducible by anhydrotetracycline (aTc). pZE12S# and pZA31R# plasmids, harboring specific combinations of RyhB mutant and *sodB* mutant, respectively, were subsequently transformed into *E. coli* ZZS00 cells derived from K-12 BW25113, harboring constitutive expression of the regulators TetR and LacI and with the native *ryhB* gene deleted (16). In some cases we incorporated a $P_{LTet\text{-}O1}$-driven *ryhB* or its mutant chromosomally and monitored their effects on the endogenous targets of RyhB using quantitative real-time PCR. Tables S1–S3 contain a summary of all of the strains, plasmids, and primers used. The sequence fragments corresponding to the various mutants of *ryhB* and *sodB* are listed in Tables S4–S6.

**Graded Tuning of Repression.** Expression of the *sodB-gfp* fusion was quantified for each strain during exponential growth in glucose minimal medium containing varying levels of the inducers and appropriate antibiotics. First, we characterized the effect of a series of RyhB derivatives on IPTG-induced expression of the wild-type sodB-GFP. These derivatives (expressed in strains ZZS00-R1 to ZZS00-R11 and referred to as R1 to R11 for brevity), contain one to three mutations in positions corresponding to the middle hairpin of the wild-type RyhB structure (nucleotide positions 32 through 56) (Fig. S1 and Table S4). The fold-repression, defined as the ratio of sodB-GFP expressions obtained at 0 and 10 ng/mL aTc with 1 mM IPTG, is shown in Fig. 1*A* (data in Table S7) for each strain characterized. The results are clearly divided into two groups: a few derivatives (R6, R7, R9) gave ≈10-fold repression, similar to the wild-type RyhB (W), whereas the rest hardly showed any response. Inspection of the sequences (Table S4) reveals that none of the RyhB derivatives maintaining strong repression has any alteration of the core interaction region (red nucleotides in Fig. S1), whereas the ones showing no repression all have at least one substitution within the core. These results demonstrate that functional RyhB–*sodB* interaction can be disrupted by a single point substitution in the core (e.g., R1 and R5). We further verified that a single point substitution in the complementary region of *sodB* made it nonresponsive to wild-type RyhB, whereas compensating mutation restoring sequence complementarity restored repression (Fig. S2).

To determine whether complementarity of the core interaction region is sufficient for repression, we generated a second group of mutations (strains ZZS00-C1 through ZZS00-C15, referred to as C1–C15) consisting of all 15 point substitutions at the two positions immediately 5′ to the start codon of *sodB* (indicated by the black box in Fig. S1), together with the complementary mutations at the corresponding RyhB positions (sequences listed in Table S5). These two positions were chosen because they do not overlap the known functional sites of *sodB* (i.e., the upstream Shine-Dalgarno sequence and the downstream start codon), which might significantly affect the translational efficiency of the *sodB-gfp* mRNA in ways unrelated to the sRNA. Surprisingly, the C-mutants showed a continuum degree of repression (Fig. 1*B*), between 1- and 10-fold. Thus, even for perfectly complementary core sequences, significant differences in expression can be easily generated through choices of the complementary bases. This differential expression is unlikely to be due to differences in the expression levels of the RyhB mutants, because most of the characterized mutants expressed at levels within two to threefold of the wild type, without noticeable correlation to the repression effects they exerted (Fig. S3, with data in Table S8).

**Energetics of the RyhB–*sodB* Interaction.** The role that the energetics of RNA interaction may play in defining the regulatory properties of the sRNAs is not known, although the energetics is often used as a component in guiding bioinformatic searches of sRNA targets (42). To determine the extent to which the observed changes in sodB-GFP expression may be accounted for by the thermodynamics of RNA–RNA interaction, we computed the free energy of duplex formation, $\Delta E$, defined as the difference between the free energy of the RyhB-*sodB* duplex ($E_{\text{duplex}}$) and the sum of the self-binding free energy of RyhB and the *sodB* control region ($E_{\text{RyhB}}$ and $E_{\text{sodB}}$, respectively) for each mutant RyhB–*sodB* pair (*SI Materials and Methods*). The results are listed in Table S9.

We investigated possible correlations between this duplex formation energy and the measured fold-repression for the various mutants characterized. Out of the 15 mutants in the C-series, several were found to have reduced expression levels with altered

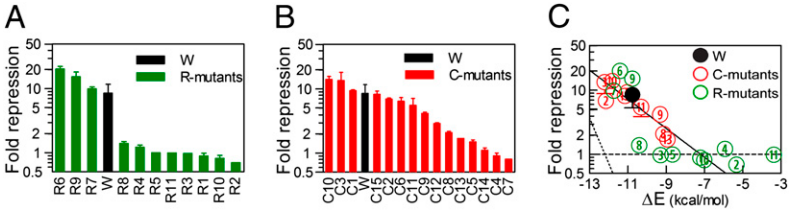


**Fig. 1.** Regulation by mutants of RyhB and *sodB*. Bar graphs showing the degrees of repression by the R-mutants (*A*) and C-mutants (*B*). (*C*) Correlation of fold-repression with the computed duplex formation energy $\Delta E$ for RyhB-*sodB* mutant pairs (red circles) and the wild type (black circle). Solid black line is the best fit of the red and black circles to the form $e^{-\beta\Delta E}$ with $\beta^{-1}$ ≈1.9 kcal/mol. Dotted black line indicates the expected correlation according to the Boltzmann distribution. Horizontal dashed line indicates the lack of correlation for the R-mutants (green circles). The duplex formation energies are listed in Table S9.

self-binding energy even in the absence of RyhB (Fig. S4, with data in Table S10). This may be attributed to alternative secondary structures of the *sodB* mRNA formed in the vicinity of the start codon (Fig. S4), and the corresponding mutants were dropped from further analysis. The remaining nine *sodB* mutants exhibited expression very similar to that of the wild type in the absence of RyhB. Their interactions with the complementary RyhB mutants were examined by plotting the duplex formation energies (ΔE) with the degrees of repression (red and black circles in Fig. 1*C*). A clear exponential correlation (solid black line) is revealed, according to the form $e^{-\beta\Delta E}$ with $\beta^{-1}\approx 1.9$ kcal/mol. In contrast, no correlation is apparent between fold-repression and the other energy parameters of the system (Fig. S5).

We next consider the R-mutants. The duplex formation energies of these RyhB mutants with the wild-type *sodB* are listed in Table S9, and the correlation of these energies with the fold-repression of sodB-GFP expression measured in Fig. 1*A* are shown as the green circles in Fig. 1*C*. The three mutants with large fold-repression all have negative ΔE values below that of the wild type, consistent with the exponential correlation observed for the C-mutants (solid black line in Fig. 1*C*). However, the other R-series mutants (which contained at least one substitution in the core interaction region, as shown in Table S4) gave no repression regardless of their duplex formation energies (dashed black line in Fig. 1*C*). Taken together, these results indicate that thermodynamic binding strength is a predictor of sRNA functionality only if sequence complementarity in the core interaction region is uninterrupted.

**Correlation of the Interaction Parameter with Interaction Energy.** For the mutant pairs showing exponential correlation between fold-repression and duplex formation energy (red circles in Fig. 1*C*), the dependence on energy is surprisingly weak compared with what would be naively expected (dotted black line in Fig. 1*C*) on the basis of the Boltzmann distribution of equilibrium thermodynamics at 37 °C. One possible cause of this discrepancy is that fold-repression does not directly reflect the strength of the sRNA–mRNA interaction. It was shown by Levine et al. (16) that gene expression regulated by noncatalytic sRNA exhibited a nonlinear, "threshold-linear" response. Specifically, for mRNA and sRNA transcribed at rates $\alpha_m$ and $\alpha_s$, respectively, the output is expected to follow the form

$$m = \frac{1}{2\beta_m}\left[(\alpha_m - \alpha_s - \lambda) + \sqrt{(\alpha_m - \alpha_s - \lambda)^2 + 4\lambda\alpha_m}\right], \quad [1]$$

where λ is a "leakage" parameter describing the rate that the mRNA is not codegraded with the regulatory sRNA but by sRNA-independent basal mechanisms. According to thermodynamics, we expect λ to be inversely related to the sRNA–mRNA binding constant, such that $\lambda \propto e^{\Delta E/k_BT}$, where $k_BT\approx 0.62$ kcal/mol.

The parameter λ can be inferred for each mutant strain. To do so, we characterized the expression levels of sodB-GFP at various IPTG levels with aTc at either 0 or 10 ng/mL for a number of C-mutants (Fig. 2*A*). Following the analysis of ref. 16, we took the measured GFP expression with/without RyhB expression to be proportional to $m$ and $\alpha_m$ respectively. We then fitted the expression data to Eq. **1**, using a single parameter $\alpha_s$ (characterizing the degree of sRNA expression) for all of the strains and a strain-dependent λ (Table S11). The best-fit curves shown in Fig. 2*A* describe the expression data well. Plotting the deduced values of λ with the duplex formation energies $\Delta E$ again reveals an exponential correlation, $\lambda \propto e^{\beta\Delta E}$, with $\beta^{-1}\approx 1.3$ kcal/mol (solid black line in Fig. 2*B*). The result is, however, still substantially different from the thermodynamic expectation (dotted black line).

**Effect of the Hfq-Binding Region.** We next examined the effect of mutated sequences in the Hfq-binding region on the function of RyhB. Hfq is required for gene silencing by RyhB (6, 35), and the AU-rich linker region that Hfq binds to has been used as a cue in the bioinformatic search of sRNA genes (43, 44). We generated the H-series mutants (H1–H19) by varying the 12 bases at positions 57–68 of RyhB, which are shown as the blue bases between the second and third hairpin in Fig. S1; the sequences are given in Table S6. The expression levels of a number of H-mutants were characterized and found comparable to the wild-type RyhB (Fig. S3). The effects of these mutants on the expression of the wild-type sodB-GFP reporter were characterized next. Because the mutated linker region is involved in RyhB–Hfq interaction but away from the region where wild-type RyhB interacts with its targets, one might expect differences in sodB-GFP expression to reflect primarily functional effects of the RyhB–Hfq interaction, including the known effect of Hfq on RyhB stability (32, 34) and possibly also the proposed effect of Hfq on RyhB–*sodB* interaction (35). As shown in Fig. 3*A*, the H-mutants exhibited >10-fold difference in their abilities to repress the target *sodB-gfp* (at 1 mM IPTG induction). Fold-repression for most of the mutants correlated exponentially with the energy cost ($\Delta E_{\text{linker}}$) of keeping the linker region open (data in Table S12, *SI Materials and Methods* gives the definition and calculation of $\Delta E_{\text{linker}}$). This finding is consistent with the expectation that RyhB function requires the binding of Hfq to the linker region, because the accessibility of the linker region (and hence the association of the Hfq to the linker) is expected to be an exponentially decreasing function of $\Delta E_{\text{linker}}$ according to thermodynamics.

It is interesting to examine the mutants deviating from the exponential correlation. In principle, with enough mutations in the linker region the secondary structure of the molecule can be significantly altered, making its function uncorrelated to the accessibility of the linker. The effect is likely to reduce the activity of the mutant, because the altered structure may not have its interaction region exposed; additionally, unprotected RNAs are known to degrade rapidly (32, 34). It is therefore rather intriguing to find a mutant (H11) that repressed sodB-GFP even more strongly than the wild type, despite a rather high linker opening cost. This mutant (RyhB-rh11) has five substitutions in the linker region (Table S6). Inspection of its minimal free energy structure (Fig. 3*B*) reveals that indeed its secondary structure may be significantly altered, with the linker region predicted to base pair with the first hairpin (positions 21–30) of the wild-type RyhB structure (Fig. S1). Interestingly, the interaction region of this mutant (red bases in Fig. 3*B*) remains open in the minimal free energy structure despite significant rearrangement of the structure elsewhere. The high abundance of the H11 mutant (Fig. S3) and its ability to repress target expression raise questions regarding the necessity of an accessible linker region.

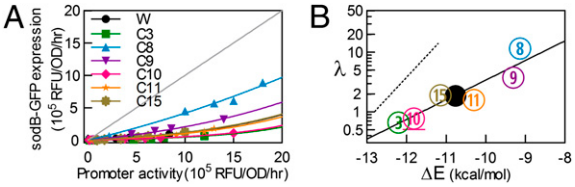


**Fig. 2.** Repression characteristics by RyhB and the energetics of RyhB–*sodB* interaction. (*A*) sodB-GFP expressions in the presence and absence of RyhB are shown in the plot, for different degrees of inducer-controlled sodB-GFP expression. Data for each mutant RyhB-*sodB* pair are indicated by a different set of symbol and color. Lines of corresponding colors indicate the results of fitting to Eq. **1**. The best-fit value of $\alpha_s = 21.3$ nM/min is comparable to the results of ref. 18 at the same level of RyhB induction. Best-fit values of λ are listed in Table S11. Diagonal gray line indicates the absence of repression. (*B*) Correlation between the RyhB–*sodB* duplex formation energy $\Delta E$ and the interaction parameters λ obtained from the fits shown in *A*. Solid black line indicates the best fit of the results to the exponential form $e^{\beta\Delta E}$ with $\beta^{-1}\approx 1.3$ kcal/mol. Dotted black line indicates the expected correlation according to the Boltzmann distribution.

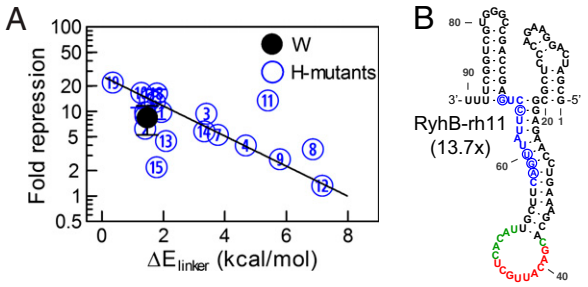


**Fig. 3.** Effects of RNA–Hfq interactions. A series of 19 RyhB mutants was constructed by random oligosynthesis. Each of these mutants has a number of bases in the linker region (positions 57–68 in Fig. S1) replaced by a different base. (*A*) The repressive effect of these mutant RyhBs on sodB-GFP expression (induced by 1 mM IPTG) is seen to correlate well with the energy cost ($\Delta E_{linker}$) of keeping the 12-base linker region open. (*B*) The minimal free energy structure of the mutant RyhB-rh11 expressed in strain ZZS00-H11. This mutant gave strong repression (13.7×) despite a large value of $\Delta E_{linker}$. The structure of this mutant is predicted to have the interaction region completely open. The bases are colored by the same color scheme as that used in Fig. S1; the substituted bases are circled.

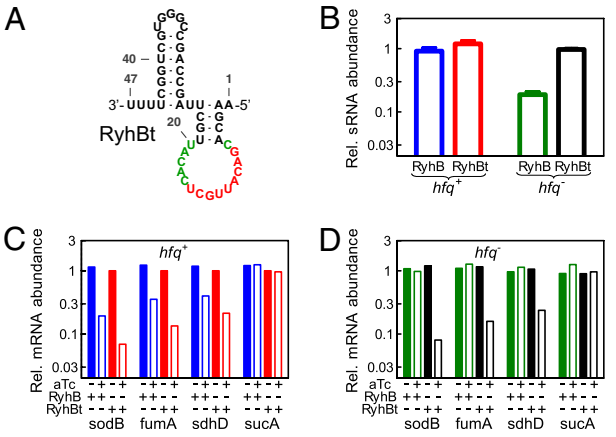


**Fig. 4.** Functional effects exerted by the truncated RyhB. (*A*) Predicted minimal free-energy structure of the mutant RyhBt, which is truncated of the first hairpin and the linker region of the wild-type RyhB (Fig. S1). (*B*) Relative expression of the chromosomally encoded RyhB and RyhBt in *hfq*$^+$ (blue and red bars) and *hfq*$^-$ background (green and black bars). (*C* and *D*) Effect of chromosomally expressed RyhB and RyhBt on the expression of various endogenous targets (*sodB, fumA, sdhD*, and *sucA*) in *hfq*$^+$ strains (*C*) and *hfq*$^-$ strains (*D*). sRNA and mRNA abundances were determined using quantitative real-time PCR in strains harboring chromosomal $P_{LTet\text{-}O1}$*:ryhB* (strain ZZS0R, ZZS0Rq, blue and green bars) or $P_{LTet\text{-}O1}$*:ryhBt* (strain ZZS0T, ZZS0Tq, red and black bars). Open and solid bars refer to results from cells grown with 10 ng/mL and no aTc, respectively. For each RNA species characterized, the abundance was normalized to the level of 16S RNA (encoded by *rrsB*), which is relatively constant in the different strains and conditions. In *C* and *D* the mRNA abundances were given relative to the mRNA level in strain ZZS0T grown with no aTc. The numerical values are listed in Tables S13, S14, and S15.

**Function of the Truncated RyhB.** We performed experiments to test the necessity of the Hfq-binding region in mediating RyhB's function. A truncated RyhB mutant (RyhBt) was constructed, removing the entire linker region (position 57–68) as well as the first hairpin (position 1–31) from the wild-type structure shown in Fig. S1. We chose this sequence because a minimal free energy structure of this sequence (Fig. 4*A*), consisting of the second hairpin (positions 32–56) of the wild-type structure with the exposed interaction region (red bases) followed immediately by the Rho-independent terminator (positions 69–90), preserves the core interaction part of RyhB. We incorporated RyhBt into the pZA31 plasmid (pZA31Rt) and characterized its effect on the expression of wild-type sodB-GFP in strain ZZS00-Rt. RyhBt is seen to exhibit strong repression, at levels comparable to the effect of wild-type RyhB in strain ZZS00-W, across the range of target expression (Fig. S6). Thus, the linker region seems to be dispensable for the repressive effect of RyhB on sodB-GFP.

To determine whether this surprising effect persists at lower levels of RyhBt expression, we constructed strains ZZS0R and ZZS0T, harboring chromosomally encoded *ryhB* and *ryhBt*, respectively, both driven by the $P_{Ltet\text{-}O1}$ promoter (Table S1). The expression levels of RyhB and RyhBt under full aTc induction are found to be comparable as characterized by quantitative real-time PCR (blue and red bars in Fig. 4*B*). We next used quantitative real-time PCR to quantify the effect of RyhBt on the expression of *sodB*, *fumA*, and *sdhD*, which are all well-established endogenous targets of RyhB (6). As shown in Fig. 4*C*, RyhBt repressed these targets 20-fold, 10-fold, and 5-fold, respectively (red bars); the degrees of repression in fact exceeded those of the wild-type RyhB (blue bars) for each target. As a negative control, neither RyhB nor RyhBt repressed the expression of *sucA*, which is in the same operon as *sdhD* but is not known to be a target of RyhB.

We further tested the role of Hfq in mediating repression by RyhB and RyhBt, using strains ZZS0Rq and ZZS0Tq, respectively, both of which contain *hfq* deletion (Table S1). As expected (6, 32, 34), RyhB exhibited no repression effect to any of the tested targets in *hfq*$^-$ background (green bars in Fig. 4*D*). However, RyhBt remains active (black bars), repressing each of the RyhB targets but not the nontarget *sucA*, at a similar level as that found in *hfq*$^+$ background (red bars in Fig. 4*C*). Quantitation of the levels of RyhB and RyhBt reveals that in *hfq*$^-$ background, RyhBt remained expressed at the same level, but the level of RyhB dropped significantly compared with that in *hfq*$^+$ background (Fig. 4*B*, black and green bars). The latter is consistent with the known instability of RyhB in *hfq*$^-$ background (32, 34), whereas the former reaffirms the observed activity of RyhBt in *hfq*$^-$ background.

## Discussion

Predictive understanding of the sequence–function relation is one of the grand challenges of systems biology. In the case of transcriptional control, a great deal is understood owing to quantitative molecular studies of protein–DNA interaction pioneered by von Hippel and collaborators (21). Even though such sequence–function relations were established only for a few exemplary systems (21–23), firm knowledge on these systems stimulated a large body of later computational work to identify and characterize TF binding sites across the genomes of organisms in the postgenome era (26).

In this work, we intended to probe a similar sequence–function relation for the interaction between sRNAs and their targets using an exemplary sRNA–mRNA system in vivo. We chose RyhB and *sodB* of *E. coli*, both well characterized at the molecular and biochemical levels. Three series of mutants were generated: the R-series featuring mutants of RyhB in the vicinity of the core interaction region, the C-series featuring mutated but complementary sequences of RyhB and *sodB-gfp* translational fusion within the core interaction region, and the H-series featuring RyhB mutants in the Hfq-binding linker region (Fig. S1). For each mutant series, we characterized the expressions of a selected number of mutants (Fig. S3) and quantified the activity of each mutant on target expression using sodB-GFP reporters. The gene expression levels were then correlated to various energetic characteristics calculated according to the available RNA-folding algorithms.

**Sequence Dependence of sRNA Function.** The R-mutants showed that one or more mismatches in the core interaction region of RyhB resulted in the complete loss of its repression on sodB-GFP expression, whereas point substitutions in the immediate

vicinity of the core may have no effect on repression (Fig. 1*A*). The lost function due to mutation in the core of either RyhB or *sodB* can be restored by complementary changes in the target sequence (Fig. S2). These findings reinforce earlier results on the SgrS-*ptsG* system by Aiba and collaborators (13, 39). In the present study, we find additionally that when the core sequence is perturbed, sRNA–mRNA binding energy is not a good predictor of functionality (green symbols in Fig. 1*C*).

The C-mutants surprisingly exhibited a continuum of repression. It is generally believed that a complementary core along with other supplemental features of the sRNA (e.g., a hairpin and an Hfq binding region) is sufficient to silence its target mRNA (1). Our results suggest instead that the identities of the complementary core nucleotides can be used to tune the strength of sRNA–mRNA interaction over a relevant range. For example, strain C8, containing a single pair of changes (C:G to A:U) compared with the wild type, exhibited only a twofold repression compared with 8.5-fold repression by the wild type.

**Energetics of RNA–RNA Interaction.** The difference between C:G and A:U pairing suggests a role exerted by the pairing energy between RyhB and *sodB*. The computed free energy of the RyhB-*sodB* duplex formation, a key intermediate in RyhB-mediated repression (32, 35), exhibits a clear exponential correlation with the fold-repression of the corresponding RyhB-*sodB* pair (Fig. 1*C*, red circles and solid black line). The exponential dependence is qualitatively consistent with the expectation that the RyhB–sodB interaction is dominated by the thermodynamics of RNA binding. However, the slope of the solid black line describing the exponential dependence is much smaller than that expected of the Boltzmann distribution at 37 °C (Fig. 1*C*, dotted black line). To further characterize the energy dependence, gene expression was characterized more quantitatively for a number of the C-mutants to quantify the interaction parameters (Fig. 2*A*). The results (Fig. 2*B*) still exhibited an exponential correlation (solid black line) and still deviated significantly from the Boltzmann distribution (dotted black line), which would be expected for an interaction driven by the thermodynamics of base pairing.

The quantitative discrepancy between the observed energy dependence of interaction is rather surprising from the perspective of molecular biophysics but is quite reasonable from the biological perspective. According to the Boltzmann distribution, there would be a 10-fold change in interaction for every 1.5 kcal/mole increase in the duplex formation energy. This would be a very large change: the smallest energy difference between two nucleotide pairings (e.g., from A:U to U:A) already involves 1 kcal/mole difference in binding energy (45). On the other hand, with the observed correlations, the degree of repression can be tuned over the functionally relevant regime (10- to 20-fold) from the choices of multiple base pairings.

The origin of the discrepancy from the Boltzmann form of energy dependence is not understood. It could be that the RyhB–*sodB* interaction was not describable by thermodynamics in vivo, in which case, however, the existence of the exponential correlation with equilibrium energy values would be perplexing. Another possibility is that the energy values used in RNA folding calculations, obtained from in vitro experiments, were systematically overestimated. In fact, in a few studies in which the results of RNA folding calculations can be compared directly to in vivo activities studied (46–48), a systematic two- to threefold overestimation of RNA binding energies has been reported. This may account for the discrepancies we observed, because the differences in the slopes of the dotted and solid lines in Figs. 1*C* and 2*B* are also two- to threefold.

**Role of Hfq and Hfq-Binding Sequence.** The H-mutants show that gradual tuning of the degree of repression (from 1- to 20-fold) can also be realized by changing the bases in the linker region (Fig. 3*A*), away from the hairpin involved in interaction with the targets. The exponential correlation of the fold-repression with the computed energy cost of opening the linker region is consistent with the notion that the binding of Hfq to the linker region is necessary for RyhB–*sodB* interaction. Regardless of the possible causes (see below), the gradual dependence of the degree of repression on the linker sequence provides another means to fine-tune the interaction. We note that tuning of the interaction strength by base changes in the linker region may be a more evolvable strategy compared with base changes in the interaction region (C-mutants), because the latter requires changes in both the sRNA and the mRNA, possibly even changes in multiple targets if the same base-pairing region is used for different targets.

Puzzling behavior exhibited by the mutant H11 led us to construct the truncated RyhB mutant, RyhBt, whose structure is expected to contain a hairpin with the open interaction region (Fig. 4*A*). Characterization of the functional effect of RyhBt, both the plasmid and chromosomally encoded versions, led to a number of surprises: first, direct characterization of RyhB and RyhBt levels (Fig. 4*B*) establishes that the level of RyhBt, which is comparable to that of RyhB in the wild-type background, is independent of Hfq. Thus, it is the Hfq-binding region that makes the wild-type RyhB unstable in *hfq*$^-$ strain. This finding is consistent with the knowledge that the AU-rich linker sequence is also the binding target of RNaseE, which degrades RNAs (32, 34), and the notion that the binding of Hfq to this region protects the RNA from cleavage. Without this RNaseE binding sequence, apparently protection by Hfq is not necessary, at least for RyhB.

Second, the Hfq-binding region of RyhB is apparently not needed for function (Figs. 4 *C* and *D*). Even in the presence of Hfq, where both RyhB and RyhBt can function, RyhBt is found to repress the endogenous RyhB targets[†] more strongly than RyhB itself for all of the cases studied (Fig. 4*C*), even though the expression levels of RyhB and RyhBt are comparable (Fig. 4*B*). Although the dependence of wild-type RyhB's function on Hfq is well known and attributed to the higher turnover rate of RyhB in *hfq*$^-$ strain (32), it is remarkable that RyhBt repressed the endogenous RyhB targets to the same degree with or without Hfq (Fig. 4 *C* and *D*). In *Vibrio cholerae*, the sRNA VrrA was shown to reduce the level of its target OmpA in *hfq*$^-$ background when overexpressed, although the degree of repression was weaker than that in *hfq*$^+$ background (49). Recently, it was also shown that positive regulation of the *rpoS* mRNA by the sRNA DsrA occurred in the absence of Hfq when DsrA was overexpressed (50). These findings led to the views that Hfq might not be essential under high concentrations of sRNA (3, 50) or under conditions whereby the sRNA and its target could form a stable complex on their own (50). In the case of RyhBt, we see that overexpression of the sRNA is not even necessary. The dependence of sRNA-mediated gene silencing has been reported for many sRNAs studied and has been used as a defining feature of this class of sRNAs (29). Because the activity of RyhBt does not require Hfq, obvious questions are raised concerning the necessity and function of the linker region (and consequently the reliance on Hfq) in these Hfq-binding bacterial sRNAs.

The independence of RyhBt's function on Hfq has another strong implication. It has been established in vitro that Hfq modifies the secondary structure of sodB by opening up the RyhB-binding region (35), and it is commonly assumed that this effect is important for RyhB repression on its targets in vivo. Enhancement of base pairing between sRNAs and target mRNAs by Hfq has also been shown in numerous other cases of sRNA-mediated repression, including Spot42–*galK* (7), MicA–*ompA* (11), SgrS–

[†]The repression of sdhD by the truncated RyhB is worth noting. It was pointed out long ago (6) that the wild-type RyhB contained sequence with extended complementarity to the translational initiation region of *sdhD*, the second gene of the *sdhCDAB* operon. In fact, the region of RyhB complementary to the translational initiation region is contained in the first hairpin of RyhB, which is deleted from RyhBt. Our results suggest that the second hairpin, which has a 10-nt continuous match with the region ending 13 nucleotides upstream of the sdhD Shine-Dalgarno sequence, is sufficient to repress sdhD. Base pairing at a distance of 10–15 nt *downstream* of the start codon was reported to be sufficient for repression (37). Apparently, similar action could also be accomplished *upstream*, as encountered here.

*ptsG* (13), and OxyS–*fhlA* (30). Together, these results project a model in which Hfq functions as a RNA chaperone mediating sRNA–mRNA interaction. Our results show that, at least for *sodB*, *fumA*, and *sdhD*, Hfq-target interaction is not necessary for repression by RyhBt.

## Materials and Methods

**Strains and Plasmids.** We constructed a series of reporter systems to quantify the RyhB–*sodB* interaction as in ref. 16. Experiments were performed in ZZS00 cells (16). Two types of plasmids, one carrying *ryhB* or its mutant (pZA31R#) and the other carrying the translational fusion of *sodB* (or its mutant) with the reporter gene *gfpmut3b* (pZE12S#), were transformed into ZZS00 to generate three series of mutant strains ZZS00-R#, ZZS00-C#, and ZZS00-H# (Table S1). In some cases, the wild-type *ryhB* and the truncated *ryhB* (*ryhBt*) in respective pZA31R and pZA31Rt plasmids were integrated into the *ryhB* locus of the chromosome.

**Medium, Growth, Measurements.** The ZZS00 cells carrying the appropriate plasmids were grown to mid-log phase in M63 minimal media at 37 °C with 0.5% glucose and the appropriate antibiotics. The cells were diluted (1:250) to fresh media and shaken overnight. The cultures were diluted into fresh M63 media ($OD_{600}$=0.002) containing the antibiotics and carbon source, as well as varying amounts of the inducers (aTc and IPTG) in wells of 48-well plates. The plates were incubated with shaking at 37 °C and examined for $OD_{600}$ and fluorescence measurements every 0.5–1 h for up to 10 h. Each measurement was repeated three times, and the data were analyzed similarly as in ref. 16. For real-time PCR analyses, total RNA was prepared using a Qiagen RNeasy Mini-prep kit or a miRNeasy Mini Kit. RNA samples were treated with the Ambion Turbo DNA-free DNase. Either a dilution series of RNA was used, or 50 ng RNA (for target genes) and 0.5 ng RNA (for *rrsB*) were used for cDNA synthesis and amplification reaction using the Bio-Rad One-Step RT-PCR Kit. Real-time PCR was performed in the Bio-Rad iQ5 Real Time PCR System. See all details in *SI Materials and Methods*.

**ACKNOWLEDGMENTS.** T.H. thanks Han Lim for helpful suggestions on qRT-PCR, and H.S. thanks Zhongcan Ouyang, Weimou Zheng for support. This work was supported by National Institutes of Health Grant R01-GM77298 (to T.H.), National Natural Science Foundation of China Grants 10428409 and 90403140, and National Basic Research Program of China Grant 2007CB814800 (to H.S.).

1. Gottesman S (2004) The small RNA regulators of *Escherichia coli*: Roles and mechanisms. *Annu Rev Microbiol* 58:303–328.
2. Flynt AS, Lai EC (2008) Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat Rev Genet* 9:831–842.
3. Waters LS, Storz G (2009) Regulatory RNAs in bacteria. *Cell* 136:615–628.
4. Altuvia S, Weinstein-Fischer D, Zhang A, Postow L, Storz G (1997) A small, stable RNA induced by oxidative stress: Role as a pleiotropic regulator and antimutator. *Cell* 90: 43–53.
5. Majdalani N, Cunning C, Sledjeski D, Elliott T, Gottesman S (1998) DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci USA* 95:12462–12467.
6. Massé E, Gottesman S (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci USA* 99:4620–4625.
7. Møller T, Franch T, Udesen C, Gerdes K, Valentin-Hansen P (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev* 16: 1696–1706.
8. Vanderpool CK, Gottesman S (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* 54:1076–1089.
9. Lenz DH, et al. (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* 118:69–82.
10. Chen S, Zhang A, Blyn LB, Storz G (2004) MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*. *J Bacteriol* 186:6689–6697.
11. Rasmussen AA, et al. (2005) Regulation of ompA mRNA stability: The role of a small regulatory RNA in growth phase-dependent control. *Mol Microbiol* 58:1421–1429.
12. Vogel J, Papenfort K (2006) Small non-coding RNAs and the bacterial outer membrane. *Curr Opin Microbiol* 9:605–611.
13. Kawamoto H, Koide Y, Morita T, Aiba H (2006) Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol Microbiol* 61:1013–1022.
14. Gerdes K, Wagner EG (2007) RNA antitoxins. *Curr Opin Microbiol* 10:117–124.
15. Papenfort K, et al. (2009) Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol Microbiol* 74:139–158.
16. Levine E, Zhang Z, Kuhlman T, Hwa T (2007) Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* 5:e229.
17. Shimoni Y, et al. (2007) Regulation of gene expression by small non-coding RNAs: A quantitative view. *Mol Syst Biol* 3:138.
18. Levine E, Hwa T (2008) Small RNAs establish gene expression thresholds. *Curr Opin Microbiol* 11:574–579.
19. Mehta P, Goyal S, Wingreen NS (2008) A quantitative comparison of sRNA-based and protein-based gene regulation. *Mol Syst Biol* 4:221.
20. Mitarai N, et al. (2009) Dynamic features of gene expression control by small regulatory RNAs. *Proc Natl Acad Sci USA* 106:10655–10659.
21. Berg OG, von Hippel PH (1988) Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor protein to recognition sites. *J Mol Biol* 200:709–723.
22. Betz JL, Sasmor HM, Buck F, Insley MY, Caruthers MH (1986) Base substitution mutants of the lac operator: In vivo and in vitro affinities for lac repressor. *Gene* 50: 123–132.
23. Fields DS, He Y, Al-Uzri AY, Stormo GD (1997) Quantitative specificity of the Mnt repressor. *J Mol Biol* 271:178–194.
24. Gerland U, Moroz JD, Hwa T (2002) Physical constraints and functional characteristics of transcription factor-DNA interaction. *Proc Natl Acad Sci USA* 99:12015–12020.
25. Bintu L, et al. (2005) Transcriptional regulation by the numbers: Models. *Curr Opin Genet Dev* 15:116–124.
26. Sahota G, Stormo GD (2010) Novel sequence-based method for identifying transcription factor binding sites in prokaryotic genomes. *Bioinformatics* 26:2672–2677.
27. Gerland U, Hwa T (2002) On the selection and evolution of regulatory DNA motifs. *J Mol Evol* 55:386–400.
28. Mustonen V, Lässig M (2005) Evolutionary population genetics of promoters: predicting binding sites and functional phylogenies. *Proc Natl Acad Sci USA* 102: 15936–15941.
29. Aiba H (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol* 10:134–139.
30. Zhang A, Wassarman KM, Ortega J, Steven AC, Storz G (2002) The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol Cell* 9:11–22.
31. Sledjeski DD, Whitman C, Zhang A (2001) Hfq is necessary for regulation by the untranslated RNA DsrA. *J Bacteriol* 183:1997–2005.
32. Massé E, Escorcia FE, Gottesman S (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev* 17:2374–2383.
33. Sittka A, et al. (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* 4:e1000163.
34. Moll I, Afonyushkin T, Vytvytska O, Kaberdin VR, Bläsi U (2003) Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* 9: 1308–1314.
35. Geissmann TA, Touati D (2004) Hfq, a new chaperoning role: Binding to messenger RNA determines access for small RNA regulator. *EMBO J* 23:396–405.
36. Soper TJ, Woodson SA (2008) The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* 14:1907–1917.
37. Bouvier M, Sharma CM, Mika F, Nierhaus KH, Vogel J (2008) Small RNA binding to 5′ mRNA coding region inhibits translational initiation. *Mol Cell* 32:827–837.
38. Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* 16:840–846.
39. Maki K, Morita T, Otaka H, Aiba H (2010) A minimal base-pairing region of a bacterial small RNA SgrS required for translational repression of ptsG mRNA. *Mol Microbiol* 76: 782–792.
40. Brennan RG, Link TM (2007) Hfq structure, function and ligand binding. *Curr Opin Microbiol* 10:125–133.
41. Massé E, Vanderpool CK, Gottesman S (2005) Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* 187:6962–6971.
42. Backofen R, Hess WR (2010) Computational prediction of sRNAs and their targets in bacteria. *RNA Biol* 7:33–42.
43. Kulkarni RV, Kulkarni PR (2007) Computational approaches for the discovery of bacterial small RNAs. *Methods* 43:131–139.
44. Chang TH, et al. (2010) Prediction of small non-coding RNA in bacterial genomes using support vector machines. *Expert Syst Appl* 37:5549–5557.
45. Mathews DH, Burkard ME, Freier SM, Wyatt JR, Turner DH (1999) Predicting oligonucleotide affinity to nucleic acid targets. *RNA* 5:1458–1469.
46. Walton SP, Stephanopoulos GN, Yarmush ML, Roth CM (1999) Prediction of antisense oligonucleotide binding affinity to a structured RNA target. *Biotechnol Bioeng* 65: 1–9.
47. Walton SP, Stephanopoulos GN, Yarmush ML, Roth CM (2002) Thermodynamic and kinetic characterization of antisense oligodeoxynucleotide binding to a structured mRNA. *Biophys J* 82:366–377.
48. Lu ZJ, Mathews DH (2008) Fundamental differences in the equilibrium considerations for siRNA and antisense oligodeoxynucleotide design. *Nucleic Acids Res* 36: 3738–3745.
49. Song T, et al. (2008) A new *Vibrio cholerae* sRNA modulates colonization and affects release of outer membrane vesicles. *Mol Microbiol* 70:100–111.
50. Soper T, Mandin P, Majdalani N, Gottesman S, Woodson SA (2010) Positive regulation by small RNAs and the role of Hfq. *Proc Natl Acad Sci USA* 107:9602–9607.

# Supporting Information

## Hao et al. 10.1073/pnas.1100432108

### SI Materials and Methods

**Construction of Strains and Plasmids.** We constructed a series of reporter system to quantify RyhB–*sodB* interaction using a strategy introduced earlier by Levine et al. (1). Experiments were performed in ZZS00 cells derived from *Escherichia coli* K-12 BW25113 with chromosomal *ryhB* deleted and a cassette *spr-lacI-tetR* inserted at the *attB* site of the chromosome to provide constitutive expression of *lacI* and *tetR*. Then two types of plasmids, one carrying *ryhB* or its mutant (pZA31R#, pZA31RH#, pZA31RC#) and the other carrying the translational fusion of *sodB* (or its mutant) with the reporter gene *gfpmut3b* (pZE12SC#), were transformed into ZZS00 to generate three series of mutant strains ZZS00-R#, ZZS00-C#, and ZZS00-H#, as listed in Table S1. The strain containing plasmids harboring wild-type RyhB and wild-type *sodB* (pZA31R and pZE12S, respectively), called ZZS00-W here, is the same as ZZS23 used in ref. 1.

The small RNA (sRNA)-source plasmid was derived from the pZA31-luc*NB* plasmid, which contained p15A replication *ori* and was marked by chloramphenicol-resistance (2). The *luc* gene was driven by the synthetic $P_{LTet\text{-}O1}$ promoter inducible by anhydrotetracycline (aTc). The wild-type *ryhB* gene was cloned directly from *E. coli* K-12 and ligated into the NdeI/BamHI sites to replace the *luc* gene, yielding the wild-type RyhB-source plasmid (pZA31R).

The mRNA-source plasmid was derived from the pZE12G plasmid, which contained colEI replication *ori* and was marked by ampicillin resistance (2). The *gfpmut3b* structural gene on pZE12G was driven by the synthetic $P_{Llac\text{-}O1}$ promoter (2) inducible by isopropyl β-D-thiogalactoside (IPTG). The 5′ UTR from the control region of *sodB* (crsodB, from −1 to +88 relative to the transcriptional start site and including the first 11 codons of *sodB*) was cloned into the EcoRI and KpnI sites, yielding the wild-type *sodB* source plasmid (pZE12S).

The *sodB* mutants constituting the pZE12SC# series were amplified with two rounds of PCR. The first round of amplification was done with primer "sc-f" and "sc#-r" (#: 1–15), with wild-type *sodB* as template. Primers "s0-f" and "s0-r" were used in the second round of PCR using the first-round PCR products as templates. The products from the second PCR were digested with restriction enzymes EcoRI and KpnI, then inserted into the same sites of pZE12G, yielding various *sodB* mutants plasmids (pZE12SC#). Plasmids and primers used in this study are listed in Table S2 and Table S3, respectively.

The *ryhB* mutants constituting the pZA31RC# series were also amplified with two rounds of PCR using the above procedure. The primers used in the first round were "rc#-f" (#: 1–15) and "rc-r." We did not add template into the reaction solution of the first round because there were 17 complementary bases at the 3′ end of the two primers, so the templates were obtained during the amplification. Primers "r0-f" and "r0-r" were used in the second round using the products from the first round as templates. The products from the second PCR were digested using restriction enzyme NdeI and BamHI, then were inserted into the sites of pZA31-*lucNB*, yielding various *ryhB* mutant plasmids (pZA31RC#).

The truncated *ryhB* (*ryhBt*) was constructed by annealing oligonucleotides directly. The "sensechain"(AAACATATGAAGCACGACATTGCTCACATTGCTTAGCCAGCCGGGTGCTGGCTTTTTTTTTGGATCCTTT, with NdeI sites and BamHI sites underlined) and "antisensechain"(AAAGGATCCAAAAAAAAAGCCAGCACCCGGCTGGCTAAGCAATGTGAGCAATGTCGTGCTTCATATGTTT, with BamHI sites and NdeI sites underlined) were resuspended at the same molar concentration of 2 OD/100 μL in "annealing buffer" [10 mM Tris (pH 8.0), 50 mM NaCl, 1 mM EDTA]. The solutions were mixed with equal volumes in a 1.5-mL tube to be placed at 94 °C for 5 min, then the tube was slowly cooled down to room temperature (below 25 °C, for half an hour). After being stored at 4 °C for half an hour, the products were digested using NdeI and BamHI before insertion into digested pZA31-*luc*NB.

The DNA fragments of the mutant *ryhB* in the R- and H-series were synthesized using an ABI 391 DNA synthesizer following a doped oligosynthesis procedure to generate random substitution. We replaced the four reservoirs each containing a single phosphoramidite with those containing combinations of phosphoramidite (the ratio of the four different phosphoramidite was 70:10:10:10 for the H-mutants to simulate a 30% substitution frequency, and was 90:3.3:3.3:3.3 for the R-mutants to simulate a 10% substitution frequency) (3). The synthesized fragments were amplified using primers "r0-f" and "r0-r" and digested using NdeI and BamHI. They were inserted into the same sites of pZA31-*lucNB* and then were transferred into ZZS20. Then the GFP expression of these strains was characterized upon induction with 1 mM IPTG and 0 or 10 ng/mL aTc. Most of the R-mutants lost their ability to repress sodB-GFP expression, whereas the H-mutants were still able to repress to various degrees. Table S4 contains sequences for all of the R-mutants that had more than twofold repression as well as a number of randomly selected ineffective ones. For the H-mutants, all of the 19 strains tested were kept. All of the characterized strains were verified by sequencing (using primers "ZA31-f" and "ZA31-rn").

To construct a background strain ZZS00-NULL (measured as negative control), we deleted the $P_{Llac\text{-}O1}$ promoter of pZE12G to yield pNULL plasmid and then transformed both pNULL and pZA31-*lucNB* to cell ZZS00.

To determine the effects of single-copy *ryhB* or its derivatives on expression of its targets, the $P_{Ltet\text{-}O1}$-driving *ryhB* and *ryhBt* as present in respective pZA31R plasmids (Table S2) were moved to the *ryhB* locus of the ZZS00 chromosome using the method described in Klumpp et al. (4). Briefly, to make the chromosomal $P_{Ltet\text{-}O1}$ driving *ryhB*, the km:*rrnB*T:$P_{Ltet\text{-}O1}$ construct present in pKDT-*rrnB*T:$P_{Ltet\text{-}O1}$ (4) was amplified using primers PtetryhB1-P1 and PtetryhB1-P2 (Table S3). The PtetryhB1-P1 contains a 50-bp region that is homologous to the upstream region of the *ryhB* promoter, whereas PtetryhB2-P2 contains a 50-bp region that is reverse complemented to the first 50-bp region of the *ryhB* gene. The PCR products were gel purified and then electroporated into ZZS00 cells expressing the λ–Red recombinase. The cells were incubated with shaking at 37 °C for 1 h and then applied onto LB + Km agar plates. The plates were incubated at 30 °C overnight. The Km resistant colonies were verified for the substitution of the native *ryhB* promoter by colony PCR and subsequently by sequencing. The resultant strain was named ZZS0R. To make chromosomal $P_{Ltet\text{-}O1}$ driving *ryhBt*, a long reverse primer (PtetryhBt2-P2) was synthesized, which carries the entire *ryhBt* and the 24 nucleotides immediately downstream of the 9-T tract of the *ryhB* gene (Table S3). The km:*rrnB*T:$P_{Ltet\text{-}O1}$-*ryhBt* was amplified from pKDT-*rrnB*T:$P_{Ltet\text{-}O1}$ using primers PtetryhB1-P1 and PtetryhBt2-P2. The PCR products were integrated into the chromosome of ZZS00 cells as described above. The resultant strain, in which $K_m$:*rrnB*T:$P_{Ltet\text{-}O1}$-*ryhBt* is substituted for *ryhB* and its promoter, is named ZZS0T. The *hfq* mutation was transferred by P1 transduction to ZZS0R and ZZS0T, yielding strains ZZS0Rq and ZZS0Tq, respectively. All

the plasmid and chromosomal constructs were verified by PCR and DNA sequencing.

**Medium, Growth, and Measurements.** The ZZS00 cells carrying the appropriate plasmids were grown with shaking to midlog phase ($OD_{600} \approx 0.5$) in M63 minimal media at 37 °C with 0.5% glucose and standard concentrations of the appropriate antibiotics. The cells were diluted (1:250) to fresh media and shaken overnight. The cultures were diluted again into fresh M63 media ($OD_{600}$ = 0.002) containing the appropriate antibiotics and carbon source, as well as varying amounts of the inducers (aTc and IPTG) in wells of 48-well plates. The plates were incubated with shaking at 37 °C and examined for $OD_{600}$ and fluorescence measurements every 0.5–1 h for up to 10 h (until a final $OD_{600}$ of 0.2) using a Wallac Victor3 1420 multilabel counter (PerkinElmer Life Sciences). Each measurement was repeated three times, and the data were analyzed similarly as in ref. 1.

**Quantitative Real-Time PCR.** For quantitative real-time PCR, strains were cultured in liquid LB with appropriate antibiotics for 6 h. The cultures were diluted at least 1,000-fold in M63 plus 0.5% glucose and appropriate concentrations of antibiotics (Cm, Ap) and inducer (10 ng/mL aTc). After ≈13 h of growth, the cultures were inoculated at $OD_{600}$ of 0.025 to identical fresh media. When $OD_{600}$ reached ≈0.5, two samples of 0.6 mL of culture were collected and treated by RNAprotect Bacteria Reagent (Qiagen; catalog no. 76506) to inactivate RNase activities before RNA preparation. Total RNA was prepared using either a RNeasy Mini Kit (Qiagen; catalog no. 74104) or a miRNeasy Mini Kit (Qiagen; catalog no. 217004). The RNA samples were treated with Turbo DNA-free DNase (Ambion; catalog no. 1907) to remove any residual genomic DNA. Typically, 50 ng total RNA was used for cDNA synthesis and subsequent real-time PCR in the same tube using iScriptTM One-Step RT-PCR Kit with SYBR Green (Bio-Rad; catalog no. 172-8892). *rrsB*, encoding 16S RNA, was included as an internal control; in these reactions 0.5 ng total RNA was used because of the extreme abundance of *rrsB*. In some cases serial dilutions of the RNA sample were made to obtain more accurate quantification by correcting for imperfect reaction efficiency. In these cases, starting RNA was serially twofold (eight consecutive dilutions) or fivefold (five consecutive dilutions) diluted. In all cases, real-time PCR was carried out in a Bio-Rad iQ5 Multicolor Real-Time PCR Detection System. Primers used in quantitative real-time PCR are listed in Table S3.

For the data analyses, for each target the mRNA level in the $hfq^+$ strain carrying the chromosomal *ryhBt* and growing with no aTc was arbitrarily set to be 1. The mRNA levels for the same target in $hfq^+$ or $hfq^-$ strains and growing with or without aTc were shown relative to 1.

For the dilution series analysis, a line was fit to the plot of relative initial RNA amount vs. $C_t$, the cycle threshold, for both the target and reference genes. The slope of this line was used to estimate the efficiency of PCR amplification for each amplicon. The ratio of target to reference gene was then computed for each point in the dilution series using the corrected efficiencies, and the average of those values is reported as the expression level of the RNA.

**Energy Calculation and RNA Structure.** We calculated the ensemble free energy of every RNA/RNA-duplex structure with or without constraint using the Vienna RNA Package (http://www.tbi.univie.ac.at/RNA/). $E_{RyhB}$, $E_{sodB}$, and $E_{duplex}$, which denote the self-binding free energy for RyhB, self-binding free energy for the *sodB* control region, and the free energy of RyhB-*sodB* duplex, respectively, were obtained without constraint. For wild-type *sodB* and its mutants, we forced both the interaction core region (52–60 in Fig. S1) and Hfq-binding site (29–44 in Fig. S1) single stranded, and calculated the difference ($E^*_{sodB}$) between the free energy of the constrained and unconstrained structures. $\Delta E_{linker}$ for wild-type RyhB and its mutants were calculated as the difference between the free energy of unconstrained structure and the structure constrained Hfq-binding sites (positions 57–68 in Fig. S1) single stranded. The minimal free energy structures were predicted by the RNAstructure software (http://rna.urmc.rochester.edu/RNAstructure.html).

**Data Analysis.** ***GFP expression.*** The data were obtained from different repeats for each combination of strain and inducers. Following the analysis of Levine et al. (1), we first obtained the cell doubling rate [μ, the slope of a linear fit of $\log_2(OD_{600})$ vs. time] for each strain and condition, yielding a doubling time of ≈1.2 h for most strains. Next, for all of the time points, we plotted the average fluorescence vs. average $OD_{600}$ (the background fluorescence production rate, which was obtained in the same way from the negative control strain ZZS00-NULL, had been removed) and extracted the slope (f). Taking account of the maturation kinetics of GFPmut3 (maturation half-life Γ ≈30 min), we computed the raw fluorescence production rate per growing cell $f\mu(1+\mu\Gamma)$ as GFP expression (5).

***Global fit.*** To fit the experimental data with the steady-state solution (Eq. **1**, main text), we assumed that the GFP expression defined above is proportional to the steady-state mRNA level m (i.e., GFP expression = $b{\cdot}m$, where $b$ reflects the rate of GFP translation and maturation). We first measured the expression of wild-type sodB-GFP under various concentration of IPTG (0.04–1 mM) and no aTc, yielding a relationship of every unit (1 nM/min) of $\alpha_m$ amounts to GFP expression of ≈116,000RFU/OD/h.

$$\begin{aligned}\text{GFP expression} &= b \cdot m = f(\alpha_m, \alpha_s, \lambda) \\ &= \frac{1}{2\beta_m}[(\alpha_m - \alpha_s - \lambda) + \sqrt{(\alpha_m - \alpha_s - \lambda)^2 + 4\lambda\alpha_m}]. \end{aligned} \quad \textbf{[S1]}$$

In the previous work (1), the wild-type RyhB–*sodB* interaction was discussed under different RyhB expression level, therefore $\alpha_s$ took different values for different experiments. As the parameter λ, which indicated the interaction strength between wild-type RyhB and *sodB*, was independent of RyhB activity, it was chosen as a global parameter in their work. In our mutation study, λ was chosen to indicate the interaction strength of different RyhB-*sodB* pairs. To find the difference of interaction strength among strains, we fitted the expression data to Eq. **1** (main text) using standard Levenberg-Marquardt algorithm. The best-fit parameters including a single parameter $\alpha_s$ shared for all of the seven strains (ZZS00-W, ZZS00-C3, ZZS00-C8, ZZS00-C9, ZZS00-C10, ZZS00-C11, and ZZS00-C15) and strain-dependent λs were obtained at confidence level of 95%.


1. Levine E, Zhang Z, Kuhlman T, Hwa T (2007) Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* 5:e229.
2. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* 25:1203–1210.
3. ABI (2002) *PCR-Mate EP Model 391 DNA Synthesizer: User's Manual* (Applied Biosystems, Foster City, CA).
4. Klumpp S, Zhang Z, Hwa T (2009) Growth rate-dependent global effects on gene expression in bacteria. *Cell* 139:1366–1375.
5. Leveau JH, Lindow SE (2001) Predictive and interpretive simulation of green fluorescent protein expression in reporter bacteria. *J Bacteriol* 183:6752–6762.

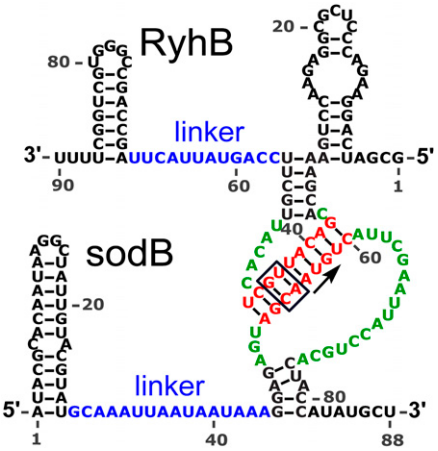


**Fig. S1.** RyhB–sodB interaction. A schematic representation of the experimentally derived interaction map between the sRNA RyhB (*Upper*) and one of its strongest targets, the mRNA of sodB (*Lower*) (1). The nucleotides in red represent the core complementarity region, which includes the start codon (AUG) of sodB (indicated by the arrow). Nonbinding nucleotides flanking the core are shown in green. The AU-rich regions (indicated in blue) bind to Hfq. Three mutant series were studied: the R-mutants contained various mutations in the interaction region of RyhB (R#), from nucleotide 32 through 56. The C-mutants include all 15 combinations of complementary point mutations of the two base pairs indicated by the black box (positions 54, 55 of sodB and 43, 44 of RyhB). The H-mutants include various mutations in the Hfq-binding region of RyhB (positions 57–68). All sequences are given in Tables S4, S5, and S6.

1. Geissmann TA, Touati D (2004) Hfq, a new chaperoning role: Binding to messenger RNA determines access for small RNA regulator. *EMBO J* 23:396–405.

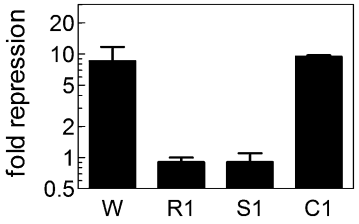


**Fig. S2.** Effect of compensatory mutations on *sodB* silencing by RyhB. Wild-type RyhB showed strong silencing ability on wild-type *sodB* (fold-repression ≈8.5). Single point substitution in RyhB (plasmid pZA31RC1, with U substituted by A at position 43 of the transcribed ryhB sequence) or *sodB* (plasmid pZE12SC1, with A substituted by U at position 55 of the transcribed sodB sequence) alone abolished repression (fold-repression ≈0.9 for both strain ZZS00-R1 and ZZS00-S1), whereas compensatory mutations restored repression (fold-repression ≈9 for strain ZZS00-C1 containing both pZA31RC1 and pZE12SC1) to a level comparable to that of the wild-type strain ZZS00-W).

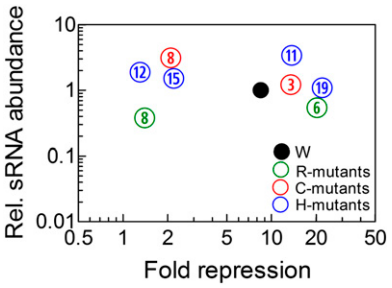


**Fig. S3.** Expression levels of plasmid-encoded *ryhB* mutants. The abundances of the wild-type RyhB (strain ZZS00-W, filled black circle) and selected mutants from the R-, C-, and H-series (green, red, and blue circles, respectively) in the absence of sodB-GFP expression were determined by quantitative real-time PCR in strains induced with 10 ng/mL aTc (*SI Materials and Methods*). The encircled number indicates the mutant number of a particular series indicated by the color (e.g., "11" in blue refers to the mutant H11). The *y* axis shows the RNA abundance of a mutant relative to the wild-type RyhB level in ZZS00-W, with the numerical values listed in Table S8. The *x* axis shows the degree of repression exerted by this mutant on sodB-GFP expression (data from Table S7). No correlation is seen between sRNA abundance and fold-repression. Note that the abundances of the RyhB mutants were mostly within two- to threefold of that of the wild type. This difference is not significant because the variation in the repeatability of these results is no less than twofold. One strain characterized (R11) did not give any RNA reading and was deemed not expressed.

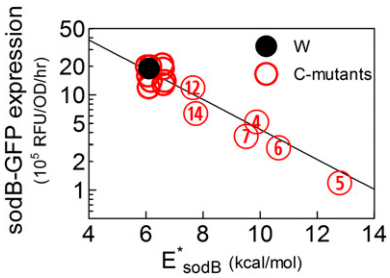


**Fig. S4.** Correlations between E*sodB and sodB-GFP expression of C-mutants in the absence of RyhB expression. We forced both the interaction core region (52–60 in Fig. S1) and Hfq-binding site (29–44 in Fig. S1) of sodB mutants single stranded, and calculated the difference (E*sodB) between the free energy of the constrained and unconstrained structures (*SI Materials and Methods*). The results are listed in Table S10. A clear exponential correlation is seen between E*sodB and sodB-GFP expression in the absence of RyhB expression, suggesting that altered sodB mRNA secondary structure in the vicinity of the start codon was responsible for the reduced expression levels observed. The black line shows the form $e^{-\beta E^*_{sodB}}$ with $\beta^{-1} \approx 2.78$ kcal/mol. The thermodynamic model has been used in explaining the control of translation by local secondary structure of mRNA in the vicinity of the start codon (1, 2), and a similar exponential relationship was detected in designing synthetic ribosome binding sites to control the expression of the red fluorescent protein (2).

1. de Smit MH, van Duin J (1994) Control of translation by mRNA secondary structure in *Escherichia coli*. A quantitative analysis of literature data. *J Mol Biol* 244:144–150.
2. Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* 27:946–950.

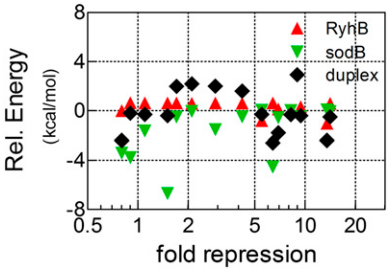


**Fig. S5.** Lack of correlation between fold-repression and the other energy scales of the system. Namely, the self-binding energies of RyhB (red) and *cr-sodB* (control region of *sodB* mRNA, green), and the RyhB-*sodB* duplex formation energy (black). The energy values here are the relative ones compared with the respective values for the wild type.

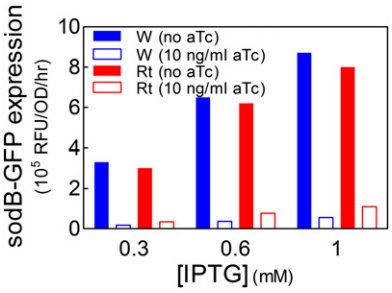


**Fig. S6.** sodB-GFP expressions in strains ZZS00-Rt and ZZS00-W with and without the inducer aTc. Red and blue bars refer to effects by RyhBt and the wild-type RyhB, respectively; open and solid bars refer to results with and without aTc, respectively.

**Table S1. Bacterial strains used in this study**

| Strain/strain series | Genotype/plasmid | Derived from | Comments |
|---|---|---|---|
| ZZS00 (1) | *ΔryhB* | BW-RI (1) | *spr-lacI-tetR* cassette derived from DH5α-ZI, *ryhB* deletion from -54 to +94 |
| ZZS20 (1) | pZE12S | ZZS00 | Wild-type *sodB* |
| ZZS00-W | pZE12S<br>pZA31R | ZZS00 | Same as ZZS23 (1) |
| ZZS00-NULL | PNULL<br>pZA31-*lucNB* (1) | ZZS00 | Negative control |
| ZZS00-S1 | pZE12SC1<br>pZA31R | ZZS00 | Mutant *sodB* (S1) with wild-type RyhB |
| ZZS00-Rt | pZE12S<br>pZA31Rt | ZZS00 | Truncated RyhB with wild-type *sodB* |
| ZZS00-R# | pZE12S<br>pZA31R# | ZZS00 | #: 1–11; muant RyhB in the core and flanking region. |
| ZZS00-H# | pZE12S<br>pZA31RH# | ZZS00 | #: 1–19; mutant RyhB in the linker region. |
| ZZS00-C# | pZE12SC#<br>pZA31RC# | ZZS00 | #: 1–15; complementary pairing of mutant RyhB and mutant *sodB* in the core region |
| ZZS0R | — | ZZS00 | $P_{Ltet\text{-}O1}$ driving *ryhB* at the *ryhB* locus of the chromosome |
| ZZS0T | — | ZZS00 | $P_{Ltet\text{-}O1}$ driving *ryhBt* at the *ryhB* locus of the chromosome |
| ZZS0Rq | — | ZZS0R | *hfq* mutation in ZZS0R |
| ZZS0Tq | — | ZZS0T | *hfq* mutation in ZZS0T |

1. Levine E, Zhang Z, Kuhlman T, Hwa T (2007) Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* 5:e229.

**Table S2. Bacterial plasmids used in this study**

| Plasmid | Genotype | Derived from | Comments |
|---|---|---|---|
| pNULL (1) | *pNULL;gfpmut3b* | pZE12 (2) | colE1 *ori*, Amp marker<br>promoter-less *gfpmut3b* |
| pZA31-*lucNB* (1) | $p_{LTet\text{-}o1}$*:luc* | pZA31-*luc* (2) | p15A *ori*, Cm marker<br>*luc* gene is flanked by NdeI site and BamHI site |
| pZE12G (1) | $P_{lac\text{-}o1}$*:gfpmut3b* | pZE12 (2) | colE1 *ori*, Amp marker<br>$p_{Llac\text{-}o1}$*: gfpmut3b* |
| pZA31R (1) | $p_{LTet\text{-}o1}$*:ryhB* | pZA31-*lucNB* | Wild-type *ryhB* |
| pZE12S (1) | $p_{Llac\text{-}o1}$*: crsodB-gfpmut3b* | pZE12G | Control region of wild-type *sodB* fused with a GFP reporter gene |
| pZA31Rt | $p_{LTet\text{-}o1}$*:ryhBt* | pZA31-*lucNB* | Truncated *ryhB* |
| pZA31R# | #: 1–11 | pZA31-*lucNB* | *ryhB* mutants r# that contain 1–3 mutations in position 32 through 56. |
| pZA31RH# | #: 1–19 | pZA31-*lucNB* | *ryhB* mutants *rh#* that contain mutations in position 57 through 68 |
| pZE12SC# | #: 1–15 | pZE12G | *sodB* mutants *sc#* with the two positions immediately 5′ to the start codon mutated. |
| pZA31RC# | #: 1–15 | pZA31-*lucNB* | *ryhB* mutants *rc#* with the complementary mutations of sc#. |

1. Levine E, Zhang Z, Kuhlman T, Hwa T (2007) Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* 5:e229.
2. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* 25: 1203–1210.

**Table S3. Primers used in this study**

| Primer | Sequence (5′ to 3′) | Comments |
|---|---|---|
| sc-f | ATACGCACAATAAGGCTATTGTACGTATGC | — |
| sc#-r | GCATATGGTAGTGCAGGTAATTCGAATGA CATXXCTACTC | #: 1–15 XX denote all 15 point substitutions at the two positions |
| s0-f | CCGGAATTCATACGCACAATAAGGCTA | *EcoRI* sites |
| s0-r | CGGGGTACCAGCATATGGTAGTGCAG | *KpnI* sites |
| rc#-f | GATCAGGAAGACCCTCGCGGAGAACCTGA AAGCACGACATXXCTCACATTGCTTCCAGT | #: 1–15XX denote all 15 point substitutions at the two positions |
| rc-r | GCGGATCCAAAAAAAAAGCCAGCACCCGGC TGGCTAAGTAATACTGGAAGCAATGTGAG | BamHI sites |
| r0-f | GGAATTCCATATGCGATCAGGAAGACCCTCG | NdeI sites |
| r0-r | CGCGGATCCAAAAAAAAAGCCAGCACCCGGC | BamHI sites |
| ZA31-f | CGACGTCTAAGAAACCAT | Universal sense primer for DNA sequencing |
| ZA31-rn | ACCGAGCGTAGCGAGTC | Anti-sense primer for *ryhB* sequencing |
| sq reverse | TTGGGACAACTCCAGTGAAA | Anti-sense primer for *sodB* sequencing |
| PtetryhB1-P1 | GGGTAAATGTCCCTTTCAACATCATTGACTTTCAA ATGCGAGTCAAATGCAGTGTAGGCTGGAGCTGCTTC | Chromosomal $P_{Ltet\text{-}O1}$ *driving ryhB* |
| PtetryhB2-P2 | GGTGGAAAAACTGGTCGTGGACGCCTGGGAACAGC GTTCATATCACGTTGGAACCTCTTACGTGCC | Chromosomal $P_{Ltet\text{-}O1}$ *driving ryhB* |
| PtetryhBt2-P2 | TGGATAAATTGAGAACGAAAGATCAAAAAAAAAGC CAGCACCCGGCTGGCTAAGCAATGTGAGCAATGTCG TGCTTTGTGCTCAGTATCTCTATCACTG | Chromosomal $P_{Ltet\text{-}O1}$ *driving ryhBt* |
| RTsodB-F | GCAATGTCATTCGAATTACCTG | *sodB* real-time PCR |
| RTsodB-R | CTGAGCTGCGTTGTTGAATACG | *sodB* real-time PCR |
| RTfumA-F | CAGTAAGTGAGAGAACAATGTC | *fumA* real-time PCR |
| RTfumA-R | CATGAACGACGCATCATGAAAC | *fumA* real-time PCR |
| RTsdhD-F | ACTGTCGTGCTTTCACTTCTCG | *sdhD* real-time PCR |
| RTsdhD-R | TGAACACTTTGGTGAACGCAGAG | *sdhD* real-time PCR |
| RTsucA-F | TTGGACTCTTCTTACCTCTCTG | *sucA* real-time PCR |
| RTsucA-R | TTGAAGAGTAACGTGAAGCGTC | *sucA* real-time PCR |
| RTryhB-F | ATATGCGATCAGGAAGACCCTC | *ryhB* and *ryhB* mutant real-time PCR |
| RTryhB-R | AAAGCCAGCACCCGGCTGGC | *ryhB* and *ryhB* mutant real-time PCR |
| RTryhBt-F | ATATGAAGCACGACATTGCTCAC | *ryhB and ryhBt* real-time PCR |
| RTryhBt-R | AAAAGCCAGCACCCGGCTGGCTAAG | *ryhB and ryhBt* real-time PCR |
| RTrrsB-F | GCTTGCTTCTTTGCTGACGAGT | *rrsB* real-time PCR |
| RT-rrsB-R | TGAGCCGTTACCCCACCTAC | *rrsB* real-time PCR |

**Table S4. R-mutants containing one to three mutations in position 32–56 of the wild-type RyhB (r1–r11)**

| Label | mutation region on *ryhB* (32–56) |
|---|---|
| **r0** | **AAGCACGACATTGCTCACATTGCTT** |
| r1 | AAGCACGACATAGCTCACATTGCTT |
| r2 | AAGCACGACATTGGTCTCATTGCTT |
| r3 | AAGCACGAATTTGCTAACATTGCTT |
| r4 | AAGCAAGACACTGCTCATATTGCTT |
| r5 | AAGCACGACATTGATCACATTGCTT |
| r6 | AAGCACGACATTGCTGACATTGCTT |
| r7 | AAGCACGACATTGCTCACACTGCTT |
| r8 | AAGCACGACGTTGCGCACATTGCTT |
| r9 | AAGCACGACATTGCTCACATTTCTT |
| r10 | AAGCATGCCAATGCTCACATTGCTT |
| r11 | AAGCACTACATGGCTAACATTGCTT |

"r0" denotes the corresponding fragment of the wild-type RyhB. Mutation points are indicated by an underline.

**Table S5. C-mutants consisting of all 15 point substitutions at the two positions 54 and 55 of *sodB* (sc1–sc15), together with the complementary mutations at the corresponding RyhB positions (rc1–rc15)**

| Mutation region on *sodB* | | Mutation region on *ryhB* | |
|---|---|---|---|
| Label | 52–60 | Label | 38–46 |
| **s0** | **AGCAATGTC** | **r0** | **GACATTGCT** |
| sc1 | AGCTATGTC | rc1 | GACATAGCT |
| sc2 | AGCGATGTC | rc2 | GACATCGCT |
| sc3 | AGCCATGTC | rc3 | GACATGGCT |
| sc4 | AGGAATGTC | rc4 | GACATTCCT |
| sc5 | AGGTATGTC | rc5 | GACATACCT |
| sc6 | AGGGATGTC | rc6 | GACATCCCT |
| sc7 | AGGCATGTC | rc7 | GACATGCCT |
| sc8 | AGAAATGTC | rc8 | GACATTTCT |
| sc9 | AGATATGTC | rc9 | GACATATCT |
| sc10 | AGAGATGTC | rc10 | GACATCTCT |
| sc11 | AGACATGTC | rc11 | GACATGTCT |
| sc12 | AGTAATGTC | rc12 | GACATTACT |
| sc13 | AGTTATGTC | rc13 | GACATAACT |
| sc14 | AGTGATGTC | rc14 | GACATCACT |
| sc15 | AGTCATGTC | rc15 | GACATGACT |

"r0" and "s0" denote the corresponding fragments of the wild-type RyhB and *sodB*, respectively. Mutation points are indicated by an underline.

**Table S6. H-mutants were generated by varying the 12 bases at positions 57–68 of RyhB (rh1–rh19)**

| Label | Mutation region on *ryhB* (57–68) |
|---|---|
| **r0** | **CCAGTATTACTT** |
| rh1 | GATGTAATACAT |
| rh2 | GCTGTTTTACAT |
| rh3 | CCAGTATTTCTT |
| rh4 | CAAGCATTGCGC |
| rh5 | CCAGTAGTTATT |
| rh6 | CATTTAATACTA |
| rh7 | GCTGTGTTAATT |
| rh8 | CCTGTCGGCGTT |
| rh9 | CGAGCAGCGTTT |
| rh10 | CTAGTAGTACTT |
| rh11 | CAGTTATTCCTG |
| rh12 | GCGGTATTCCTG |
| rh13 | CCGTTACTACTA |
| rh14 | ACAGCCTTCCTT |
| rh15 | CCGGTATTACAT |
| rh16 | CACGACATAGTT |
| rh17 | CCAGTATTACAT |
| rh18 | GCAGTATTACTT |
| rh19 | ACGTTATTACTT |

"r0" denotes the corresponding fragment of the wild-type RyhB. Mutation points are indicated by an underline.

**Table S7. Measurement of *sodB-gfp* expression and calculation of "fold-repression"**

| | *sodB-gfp* Expression (RFU/OD/hr) [IPTG] = 1 mM | | | | |
|---|---|---|---|---|---|
| Interaction pair | [aTc] = 0 | | [aTc] = 10 ng/mL | | Fold-repression |
| W | 1.3E+06 | ±2.0E+05 | 1.6E+05 | ±5.0E+04 | 8.5 ± 3.2 |
| S1 | 1.3E+06 | ±5.3E+05 | 1.5E+06 | ±4.0E+05 | 0.9 ± 0.2 |
| C1 | 1.5E+06 | ±3.0E+05 | 1.6E+05 | ±3.0E+04 | 9.4 ± 0.3 |
| C2 | 1.6E+06 | ±2.0E+05 | 2.4E+05 | ±1.0E+05 | 6.9 ± 0.3 |
| C3 | 1.2E+06 | ±1.0E+05 | 1.1E+05 | ±4.4E+04 | 13.5 ± 4.7 |
| C4 | 6.1E+05 | ±5.0E+04 | 6.9E+05 | ±1.3E+05 | 0.9 ± 0.1 |
| C5 | 1.0E+05 | ±1.2E+04 | 6.8E+04 | ±1.1E+04 | 1.5 ± 0.1 |
| C6 | 2.9E+05 | ±8.0E+04 | 4.6E+04 | ±1.7E+04 | 6.4 ± 0.9 |
| C7 | 1.2E+06 | ±0 | 1.5E+06 | ±0 | 0.8 ± 0 |
| C8 | 1.4E+06 | ±4.0E+05 | 6.9E+05 | ±2.0E+05 | 2.1 ± 0.1 |
| C9 | 1.4E+06 | ±0 | 3.3E+05 | ±1.0E+04 | 4.2 ± 0.2 |
| C10 | 1.3E+06 | ±2.0E+05 | 9.7E+04 | ±2.3E+04 | 14.1 ± 1.6 |
| C11 | 1.6E+06 | ±3.0E+05 | 3.3E+05 | ±1.5E+05 | 5.5 ± 1.6 |
| C12 | 1.6E+06 | ±1.0E+05 | 5.4E+05 | ±6.0E+04 | 2.9 ± 0.1 |
| C13 | 9.5E+05 | ±1.5E+05 | 5.5E+05 | ±7.0E+04 | 1.7 ± 0 |
| C14 | 3.7E+05 | ±0 | 3.3E+05 | ±3.0E+04 | 1.1 ± 0.1 |
| C15 | 1.3E+06 | ±3.0E+05 | 1.6E+05 | ±3.0E+04 | 8.2 ± 1.0 |
| R1 | 1.2E+06 | ±2.2E+05 | 1.4E+06 | ±2.0E+05 | 0.9 ± 0.1 |
| R2 | 1.1E+06 | ±0 | 1.6E+06 | ±1.0E+05 | 0.7 ± 0 |
| R3 | 1.2E+06 | ±0 | 1.2E+06 | ±0 | 1.0 ± 0 |
| R4 | 1.2E+06 | ±2.0E+05 | 9.9E+05 | ±1.6E+05 | 1.2 ± 0.1 |
| R5 | 1.4E+06 | ±1.0E+05 | 1.5E+06 | ±1.0E+05 | 1.0 ± 0 |
| R6 | 1.3E+06 | ±2.0E+05 | 6.4E+04 | ±5.0E+03 | 20.3 ± 2.0 |
| R7 | 1.2E+06 | ±1.0E+05 | 1.2E+05 | ±0 | 9.9 ± 0.7 |
| R8 | 1.3E+06 | ±1.0E+05 | 9.6E+05 | ±5.0E+04 | 1.4 ± 0.1 |
| R9 | 1.2E+06 | ±3.0E+05 | 7.9E+04 | ±3.1E+04 | 15.4 ± 2.8 |
| R10 | 1.5E+06 | ±1.0E+05 | 1.9E+06 | ±1.0E+05 | 0.8 ± 0.1 |
| R11 | 1.6E+06 | ±0 | 1.6E+06 | ±0 | 1.0 ± 0 |
| H1 | 1.1E+06 | ±8.0E+04 | 1.1E+05 | 1.5E+04 | 10.0 ± 0.5 |
| H2 | 8.3E+05 | ±9.0E+04 | 1.3E+05 | ±2.0E+04 | 6.3 ± 0 |
| H3 | 1.4E+06 | ±0 | 1.5E+05 | ±1.0E+04 | 9.5 ± 0.7 |
| H4 | 8.2E+05 | ±3.0E+04 | 2.0E+05 | ±1.0E+04 | 4.0 ± 0.2 |
| H5 | 8.8E+05 | ±1.0E+05 | 6.5E+04 | ±1.3E+04 | 13.9 ± 1.3 |
| H6 | 1.3E+06 | ±2.0E+05 | 1.2E+05 | ±7.0E+04 | 13.1 ± 5.1 |
| H7 | 9.5E+05 | ±8.0E+04 | 1.8E+05 | ±1.0E+04 | 5.3 ± 0.2 |
| H8 | 1.4E+06 | ±2.0E+05 | 3.9E+05 | ±4.0E+04 | 3.6 ± 0.1 |
| H9 | 9.7E+05 | ±1.0E+05 | 3.5E+05 | ±4.0E+04 | 2.7 ± 0.1 |
| H10 | 1.2E+06 | ±2.6E+05 | 7.4E+04 | ±3.6E+04 | 16.8 ± 5.7 |
| H11 | 1.0E+06 | ±1.0E+05 | 7.8E+04 | ±1.3E+04 | 13.7 ± 2.7 |
| H12 | 1.1E+06 | ±2.0E+05 | 8.4E+05 | ±5.0E+04 | 1.3 ± 0.2 |
| H13 | 1.3E+06 | ±1.0E+05 | 2.9E+05 | ±7.0E+04 | 4.5 ± 0.8 |
| H14 | 9.2E+05 | ±3.8E+05 | 1.7E+05 | ±8.1E+04 | 5.9 ± 0.7 |
| H15 | 1.2E+06 | ±0 | 5.4E+05 | ±4.0E+04 | 2.2 ± 0.1 |
| H16 | 8.7E+05 | ±3.3E+05 | 1.0E+05 | ±4.0E+04 | 8.9 ± 0.5 |
| H17 | 1.3E+06 | ±1.0E+05 | 1.3E+05 | ±2.0E+04 | 9.9 ± 0.8 |
| H18 | 1.2E+06 | ±2.0E+05 | 8.0E+04 | ±3.0E+04 | 16.5 ± 4.0 |
| H19 | 1.2E+06 | ±1.0E+05 | 5.7E+04 | ±1.4E+04 | 22.0 ± 5.2 |

Error bars were calculated according to two or more repeated measurements.

**Table S8. Quantitative real-time PCR results of expression levels of plasmid-harboring *ryhB* and *ryhB* mutants**

| Strain (ZZS00-) | Relative abundance of sRNA [aTc] = 10 ng/mL |
|---|---|
| W | 1.0 |
| R6 | 0.5 |
| R8 | 0.4 |
| R11 | 0.0001 |
| C3 | 1.2 |
| C8 | 3.2 |
| H11 | 3.4 |
| H12 | 1.9 |
| H15 | 1.5 |
| H19 | 1.1 |

Abundances of RyhB (W) and RyhB mutants (R-, C-, and H-mutants) were determined by quantitative real-time PCR in strains induced with 10 ng/mL aTc, with the level of 16S RNA as internal control; detailed description in *SI Materials and Methods*. Here we show the RNA abundances relative to the wild-type RyhB level in strain ZZS00-W.

**Table S9. Ensemble free energy predicted for each interaction pair: free energy of RyhB-*sodB* duplex ($E_{duplex}$), self-binding free energy for RyhB ($E_{RyhB}$) and control region of *sodB* ($E_{sodB}$), and free energy of duplex formation ($\Delta E$)**

| | Energy values (kcal/mol) | | | |
|---|---|---|---|---|
| Interaction pair | $E_{RyhB}$ | $E_{sodB}$ | $E_{duplex}$ | $\Delta E$ |
| W | −26.5 | −17.32 | −54.58 | −10.76 |
| C1 | −26.17 | −17.78 | −54.97 | −11.02 |
| C2 | −26.36 | −17.87 | −56.37 | −12.14 |
| C3 | −27.53 | −17.24 | −56.97 | −12.2 |
| C4 | −25.85 | −21.09 | −54.78 | −7.84 |
| C5 | −25.86 | −24.01 | −54.98 | −5.11 |
| C6 | −25.86 | −21.86 | −57.17 | −9.45 |
| C7 | −26.49 | −20.73 | −56.97 | −9.75 |
| C8 | −25.89 | −17.36 | −52.38 | −9.13 |
| C9 | −25.88 | −17.8 | −52.98 | −9.3 |
| C10 | −25.89 | −17.38 | −55.08 | −11.81 |
| C11 | −27.31 | −17.28 | −54.88 | −10.29 |
| C12 | −25.86 | −18.86 | −52.58 | −7.86 |
| C13 | −25.86 | −17.82 | −52.59 | −8.91 |
| C14 | −25.86 | −18.96 | −54.88 | −10.06 |
| C15 | −26.46 | −17.3 | −54.9 | −11.14 |
| R1 | −26.17 | −17.32 | −50.69 | −7.2 |
| R2 | −28.48 | −17.32 | −51.12 | −5.32 |
| R3 | −26.35 | −17.32 | −52.95 | −9.28 |
| R4 | −27.52 | −17.32 | −50.76 | −5.92 |
| R5 | −25.9 | −17.32 | −51.88 | −8.66 |
| R6 | −27.04 | −17.32 | −55.76 | −11.4 |
| R7 | −26.51 | −17.32 | −55.56 | −11.73 |
| R8 | −27 | −17.32 | −54.72 | −10.4 |
| R9 | −27.5 | −17.32 | −55.58 | −10.76 |
| R10 | −28.29 | −17.32 | −52.61 | −7 |
| R11 | −29.76 | −17.32 | −50.45 | −3.37 |
| H1 | −26.98 | −17.32 | −53.85 | −9.55 |
| H2 | −26.46 | −17.32 | −53.4 | −9.62 |
| H3 | −28.42 | −17.32 | −53.76 | −8.02 |
| H4 | −29.78 | −17.32 | −55.79 | −8.69 |
| H5 | −26.59 | −17.32 | −54.4 | −10.49 |
| H6 | −26.82 | −17.32 | −53.66 | −9.52 |
| H7 | −28.82 | −17.32 | −56.68 | −10.54 |
| H8 | −31.88 | −17.32 | −55.7 | −6.5 |
| H9 | −30.82 | −17.32 | −54.69 | −6.55 |
| H10 | −26.32 | −17.32 | −54.07 | −10.43 |
| H11 | −30.57 | −17.32 | −54.86 | −6.97 |
| H12 | −32.35 | −17.32 | −56.27 | −6.6 |
| H13 | −27.18 | −17.32 | −53.76 | −9.26 |
| H14 | −28.38 | −17.32 | −55.58 | −9.88 |
| H15 | −26.8 | −17.32 | −55.39 | −11.27 |
| H16 | −26.45 | −17.32 | −52.57 | −8.8 |
| H17 | −26.44 | −17.32 | −54.47 | −10.71 |
| H18 | −26.83 | −17.32 | −54.06 | −9.91 |
| H19 | −25.41 | −17.32 | −54.58 | −11.85 |

See *SI Materials and Methods*.

**Table S10. Expression levels of *sodB-gfp* in strains ZZS00-W and ZZS00-C1 through ZZS00-C15 in the absence of RyhB expression (no aTc), together with the energy cost ($E^*_{sodB}$) of opening both the interaction core region (52–60) and the Hfq-binding region (29–44) of *sodB***

| Interaction pair | *sodB-gfp* expression (RFU/OD/hr) [IPTG] = 3 mM | $E^*_{sodB}$ (kcal/mol) |
|---|---|---|
| W | 1.9E+06 | 6.11 |
| C1 | 2.1E+06 | 6.57 |
| C2 | 1.4E+06 | 6.66 |
| C3 | 2.0E+06 | 6.03 |
| C4 | 5.2E+05 | 9.88 |
| C5 | 1.2E+05 | 12.8 |
| C6 | 2.8E+05 | 10.65 |
| C7 | 3.7E+05 | 9.52 |
| C8 | 2.0E+06 | 6.15 |
| C9 | 1.9E+06 | 6.59 |
| C10 | 1.4E+06 | 6.17 |
| C11 | 1.6E+06 | 6.07 |
| C12 | 1.2E+06 | 7.65 |
| C13 | 1.3E+06 | 6.61 |
| C14 | 6.4E+05 | 7.75 |
| C15 | 1.2E+06 | 6.09 |

**Table S11. Best-fit parameters of the data in Fig. 2*A* (main text)**

| Strain | $\lambda$ (nM/min) |
|---|---|
| W | $1.85 \pm 0.48$ |
| C3 | $0.67 \pm 0.22$ |
| C8 | $11.69 \pm 2.68$ |
| C9 | $3.80 \pm 0.94$ |
| C10 | $0.78 \pm 0.28$ |
| C11 | $1.59 \pm 0.43$ |
| C15 | $1.92 \pm 0.48$ |

Best-fit value for $\alpha_s$ is $21.30 \pm 3.24$ nM/min. Detailed description in *SI Materials and Methods*.

**Table S12. AU content of the Hfq-binding region of RyhB (position 57–68) and the energy cost ($\Delta E_{linker}$) of keeping this region open**

| RyhB | No. AU | $\Delta E_{linker}$ (kcal/mol) |
|---|---|---|
| RyhB-r0 | 8 | 1.48 |
| RyhB-rh1 | 9 | 1.94 |
| RyhB-rh2 | 8 | 1.42 |
| RyhB-rh3 | 8 | 3.4 |
| RyhB-rh4 | 5 | 4.69 |
| RyhB-rh5 | 8 | 1.57 |
| RyhB-rh6 | 10 | 1.73 |
| RyhB-rh7 | 8 | 3.78 |
| RyhB-rh8 | 4 | 6.86 |
| RyhB-rh9 | 5 | 5.8 |
| RyhB-rh10 | 8 | 1.3 |
| RyhB-rh11 | 7 | 5.4 |
| RyhB-rh12 | 5 | 7.17 |
| RyhB-rh13 | 7 | 2.09 |
| RyhB-rh14 | 6 | 3.34 |
| RyhB-rh15 | 7 | 1.78 |
| RyhB-rh16 | 7 | 1.43 |
| RyhB-rh17 | 8 | 1.42 |
| RyhB-rh18 | 8 | 1.79 |
| RyhB-rh19 | 9 | 0.37 |

Definition and calculation of $\Delta E_{linker}$ values in *SI Materials and Methods*.

**Table S13. Quantitative real-time PCR results for expression of chromosomal encoded *ryhB* and *ryhBt* in the *hfq*$^+$ (ZZS0R and ZZS0T, respectively) and *hfq*$^-$ (ZZS0Rq and ZZS0Tq, respectively) strains induced with 10 ng/mL aTc**

| Strain | Relative abundance of sRNA [aTc] = 10 ng/mL |
|---|---|
| ZZS0R | 0.9 |
| ZZS0T | 1.2 |
| ZZS0Rq | 0.2 |
| ZZS0Tq | 1.0 |

RNA abundances were normalized to the level of 16S RNA (encoded by *rrsB*), which was chosen as an internal control. Detailed description in *SI Materials and Methods*.

**Table S14. Quantitative real-time PCR results for effect of chromosomal encoded *ryhB* and *ryhBt* on expression of various chromosomal targets (*sodB*, *fumA*, *sdhD*, and *sucA*) in *hfq*$^+$ strains**

| | | Relative abundance | | | |
|---|---|---|---|---|---|
| Strain | Condition | *sodB* | *fumA* | *sdhD* | *sucA* |
| ZZS0R | [aTc] = 0 | 1.1 | 1.2 | 1.2 | 1.2 |
| | [aTc] = 10 ng/mL | 0.2 | 0.4 | 0.4 | 1.2 |
| ZZS0T | [aTc] = 0 | 1.0 | 1.0 | 1.0 | 1.0 |
| | [aTc] = 10 ng/mL | 0.1 | 0.1 | 0.2 | 1.0 |

RNA abundances were normalized to the level of 16S RNA (encoded by *rrsB*), which was chosen as an internal control. Detailed description in *SI Materials and Methods*.

**Table S15. Quantitative real-time PCR results for effect of chromosomal encoded *ryhB* and *ryhBt* on expression of various chromosomal targets (*sodB*, *fumA*, *sdhD*, and *sucA*) in *hfq*$^{-}$ strains**

| Strain | Conditions | Relative abundance | | | |
|---|---|---|---|---|---|
| | | *sodB* | *fumA* | *sdhD* | *sucA* |
| ZZS0Rq | [aTc] = 0 | 1.1 | 1.1 | 1.0 | 0.9 |
| | [aTc] = 10 ng/mL | 1.0 | 1.3 | 1.1 | 1.3 |
| ZZS0Tq | [aTc] = 0 | 1.2 | 1.2 | 1.1 | 0.9 |
| | [aTc] = 10 ng/mL | 0.1 | 0.2 | 0.2 | 1.0 |

RNA abundances were normalized to the level of 16S RNA (encoded by *rrsB*), which was chosen as an internal control. Detailed description in *SI Materials and Methods*.